Belief hierarchies in standard state space models and epistemic equivalence of belief spaces

Maastricht University

Elias Tsakas

Belief hierarchies in standard state space models and epistemic equivalence of belief spaces

RM/10/048

METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Belief hierarchies in standard state space models and epistemic equivalence of belief spaces*

ELIAS TSAKAS†

Department of Economics, Maastricht University

October 7, 2010

Abstract

In this paper we formalize the notion of lexicographic belief hierarchies in standard partitional models (Aumann, 1976). We introduce the notion of epistemic equivalence between two belief space, and show that the state space representation of lexicographic belief hierarchies is equivalent to the usual type-space approach, even when the latter induces lexicographic belief hierarchies which violate mutual singularity.

KEYWORDS: Standard state space models, lexicographic belief hierarchies, epistemic equivalence.

JEL CLASSIFICATION: C70, D80, D81, D82.

1. Introduction

A belief hierarchy describes the individual's beliefs, beliefs about others' beliefs, and so on. Belief hierarchies are an integral part of modern economic theory, often used for analyzing games with incomplete information (Harsanyi, 1967-68), as well as for providing epistemic characterizations for different solution concepts, such as, for instance, rationalizability (Bernheim, 1984;

*I am extremely indebted to Amanda Friedenberg for numerous conversations on this paper. I would also like to thank Christian Bach, Erik Balder, Adam Brandenburger, Matthew Embrey, Dimitris Georgiou, Willemien Kets, Friederike Mengel, Andrés Perea, Jeff Steif, Mark Voorneveld and the seminar participants at Maastricht University for very useful discussions and comments. I would also like to thank the Olin School of Business, at Washington University in St. Louis for its hospitality while working on the first draft of this paper. Financial support from the Adlerbertska Research Foundation and the Marie Curie Fellowship is gratefully acknowledged.

†P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Tel: +31-43-38 83649; Fax: +31-43-38 84878; E-mail: e.tsakas@maastrichtuniversity.nl

Pearce, 1984), Nash equilibrium (Aumann and Brandenburger, 1995), and correlated equilibrium (Aumann, 1987).

Usual belief hierarchies are not sufficiently rich for characterizing other solution concepts, such as iterated admissibility (Brandenburger et al., 2008), or perfect equilibrium (Selten, 1975), where secondary theories (assessments) given zero probability events become crucial. Suppose, for instance, that Ann assigns probability 1 to Bob playing b_1 in a normal form game, without at the same time ruling out the possibility that he chooses b_2 or b_3; she simply deems these choices infinitely less likely than b_1. This notion is not captured by usual belief hierarchies, which do not describe Ann's conditional beliefs given the (primarily) null event "b_2 or b_3". Therefore, we need to add more structure to Ann's beliefs in order to capture the idea that Ann considers the event "Bob playing b_2" more likely than "Bob playing b_3", in case he does not end up playing b_1. This is typically done with the use of lexicographic beliefs, which enrich Ann's state of mind with a sequence of (secondary, tertiary, and so on) assessments, instead of a unique measure[1]. The primary assessment, else called theory, coincides with the standard beliefs, which assign probability 1 to b_1, whereas the secondary theory can be thought as the beliefs Ann would hold in case she was informed that Bob did not play b_1.

This idea is naturally extended to an interactive setting, by generalizing belief hierarchies to lexicographic belief hierarchies, which capture the idea that Ann's k-th order beliefs are lexicographic over the space of Bob's $(k-1)$-th order lexicographic beliefs (Brandenburger et al., 2008). Lexicographic belief hierarchies are typically represented by a generalized type space model which maps every individual type to a lexicographic probability system over the product of the fundamental space of uncertainty and the opponent's types, thus extending Harsanyi's seminal structure to lexicographic beliefs.

In this paper we provide an alternative representation of lexicographic belief hierarchies using Aumann's partitional model of differential information (Aumann, 1976). The standard requirement, whenever we obtain two different representations of the same object – lexicographic belief hierarchies in this case – is to make sure that the two are equivalent, and therefore interchangeable. It turns out that the state space representation that we introduce, satisfies our epistemic equivalence criterion, i.e., there is always a bijection between information sets (in the partitional model) and types (in the type space model) that preserves lexicographic belief hierarchies, even when mutual singularity is violated in the type space model.

The need for establishing invariance between different epistemic models has already been recognized in the literature: Brandenburger and Friedenberg (2010, p. 804) point out the equivalence

[1]For an axiomatic foundation of lexicographic beliefs, see Blume et al. (1991a).

between lexicographic and conditional probability systems[2], which is formally shown by Brandenburger et al. (2007) in a single-individual environment. They also discuss the importance of exploring whether epistemically invariant models yield the same predictions. Similar exercises have been carried out for standard belief hierarchies, by Brandenburger and Dekel (1993) who mapped Harsanyi type space to epistemically equivalent standard space models, and Tan and Werlang (1992) who constructed the converse mapping.

The paper is structured as follows: Section 2 presents some preliminaries on Polish spaces, lexicographic probability systems and conditional probability systems. In Section 3, we provide the state space representation of lexicographic belief hierarchies, and prove our equivalence results). Section 5 concludes.

2. Preliminaries

2.1. Polish spaces

We present some preliminaries on Polish spaces. For further reference see Kechris (1995). A topological space $(Z, \mathcal{T})$ is called Polish if it is separable and completely metrizable. A subspace of a separable metrizable space is also separable and metrizable. Examples of Polish spaces include countable sets endowed with the discrete topology, $\mathbb{R}^n$ endowed with the usual topology and closed subsets of Polish spaces endowed with the relative topology. The countable product of Polish spaces, endowed with the product topology, is Polish. A closed subspace of a Polish space, endowed with the relative topology, is also Polish. The topological sum of a countable collection of Polish spaces, denoted by $\oplus$, is also Polish.

For any topological space Z, let $\Delta(Z)$ denote the set of all Borel probability measures, endowed with the weak topology. If Z is Polish then so is $\Delta(Z)$ (Aliprantis and Border, 1994, p. 515). For some $p \in \Delta(Z)$, let $\Gamma(p)$ denote its support, i.e., the set of all points $z \in Z$ such that every $T \in \mathcal{T}$ with $z \in T$ has positive measure: $\Gamma(p) = \{z \in Z : z \in T \in \mathcal{T} \Rightarrow p(T) > 0\}$. The support is the smallest closed subset of Z with measure equal to 1. If Z is separable and metrizable, the support is unique (Parthasarathy, 1967, pp. 27–28).

[2]Conditional probability systems (Rênyi, 1955; Battigalli and Siniscalchi, 1999) are another way of modeling beliefs given secondary conditional hypotheses. We formally define them later in the paper.

2.2. Lexicographic probability systems

Definition 2.1. A lexicographic probability system (LPS) over a (Polish) space Z is a sequence of probability measures $\tilde{p} := (p^n)_{n=1}^N$ with $N \leq \infty$. The space of all LPS's is denoted by $\tilde{\Delta}(Z)$.

The measure p^1 represents the individual's primary theory, p^2 her secondary theory, and so on. We say that the LPS is finite whenever $N < \infty$. Let $\Gamma(\tilde{p}) := \bigcup_{n=1}^N \Gamma(p^n)$ denote the support of the LPS: $\tilde{p}$ is full-support whenever $\Gamma(\tilde{p}) = Z$.

We say that $\tilde{p}$ satisfies mutual singularity whenever for every $n = 1, .., N$ there is a Borel set F_n such that $p^n(F_n) = 1$ and $p^n(F_m) = 0$ for all $m \neq n$ (Brandenburger et al., 2008, p. 321).

The concept of weak assumption[3] that we introduce here, extends (standard) confident belief to LPS's. We say that an event $E \subseteq Z$ is weakly assumed under $\tilde{p}$ whenever $p^n(E) = 1$ for all n. This notion is weaker than assumption, as introduced by Brandenburger et al. (2008) which also requires that for all open sets T with $E \cap T \neq \emptyset$, there is some n such that $p^n(E \cap T) > 0$.

Lexicographic probability systems were introduced in the literature by Blume et al. (1991a), who provided the corresponding axiomatic foundation. They also used LPS's to epistemically characterize equilibrium refinements, such as perfect equilibrium (Blume et al., 1991b).

2.3. Conditional probability systems

Let $(\Omega, \mathcal{F})$ be a measurable space, where Ω is Polish, $\mathcal{F}$ is the Borel σ-algebra, and let $\mathcal{G} \subset \mathcal{F}$ be a collection of non-empty conditioning events (not necessarily an algebra).

Definition 2.2. A conditional probability system (CPS) on $(\Omega, \mathcal{F}, \mathcal{G})$ is a function $\pi : \mathcal{F} \times \mathcal{G} \to [0, 1]$ satisfies the following properties:

(C_1) $\pi(G|G) = 1$, if $G \in \mathcal{G}$,

(C_2) $\pi(\cdot|G)$ is a probability measure over $(\Omega, \mathcal{F})$ for every $G \in \mathcal{G}$,

(C_3) $\pi(F|G) = \pi(F|E) \times \pi(E|G)$, if $F \subseteq E \subseteq G$, and $F \in \mathcal{F}$ and $G, E \in \mathcal{G}$.

Let $\mathbb{P}$ be a prior probability measure over $(\Omega, \mathcal{F})$. Then, it follows from Ω being Polish that there is always a CPS on $(\Omega, \mathcal{F}, \mathcal{G})$, which agrees with $\mathbb{P}$ whenever possible, i.e., $\pi(F|G) = \mathbb{P}(F \cap G)/\mathbb{P}(G)$ for all $G \in \mathcal{G}$ with $\mathbb{P}(G) > 0$ (Blackwell and Dubins, 1975; Blackwell and Ryll-Nardzewski, 1963; Brandenburger and Dekel, 1987).

We say that an event $E \in \mathcal{F}$ is strongly believed in $(\Omega, \mathcal{F}, \mathcal{G})$ whenever $\pi(E|G) = 1$ for all $G \in \mathcal{G}$ such that $E \cap G \neq \emptyset$ (Battigalli and Siniscalchi, 1999, 2002).

[3] Perea (2010, p. 187) uses the term *full belief* to describe this notion.

Conditional probability spaces first appeared in Rênyi (1955), and were introduced in a game theoretic framework by Battigalli and Siniscalchi (1999).

3. Lexicographic belief hierarchies

Consider the (countable) fundamental space of uncertainty Σ with typical element σ, endowed with the discrete topology. Examples of Σ include the set of payoff functions in an incomplete information game, and the set of action profiles in a normal form game. Let $I = \{a, b\}$ be the set of individuals[4], with typical elements i and j.

Consider the following sequence:

$$\begin{aligned}
\Theta_0 &:= \Sigma \\
\Theta_1 &:= \Theta_0 \times \tilde{\Delta}(\Theta_0) \\
&\vdots \\
\Theta_k &:= \Theta_{k-1} \times \tilde{\Delta}(\Theta_{k-1}) \\
&\vdots
\end{aligned}$$

A hierarchy of lexicographic beliefs is a sequence of LPS's, $(\tilde{p}_i^1, \tilde{p}_i^2, ...) \in \times_{k=0}^{\infty} \tilde{\Delta}(\Theta_k)$, where $\tilde{p}_i^k = (p_i^{k,1}, ..., p_i^{k,N}) \in \tilde{\Delta}(\Theta_{k-1})$ denotes i's k-th order lexicographic beliefs, with $p_i^{k,n} \in \Delta(\Theta_{k-1})$ being the n-th theory of the k-th order lexicographic beliefs.

As usual, we restrict our focus to hierarchies that satisfy the standard coherency requirement. We say that $(\tilde{p}_i^1, \tilde{p}_i^2, ...)$ is coherent if for all $k > 1$, $\text{marg}_{\Theta_{k-2}} \tilde{p}_i^k = \tilde{p}_i^{k-1}$, with $\text{marg}_{\Theta_{k-2}} \tilde{p}_i^k := (\text{marg}_{\Theta_{k-2}} p_i^{k,1}, ..., \text{marg}_{\Theta_{k-2}} p_i^{k,N})$ denoting the marginal lexicographic probability system. We consider hierarchies that satisfy, not only coherency, but also common weak assumption of coherency. For the time being, we have not defined coherency of the opponent as an event, and therefore common weak assumption of coherency is not formally defined. We will formally do this in the following sections.

3.1. Lexicographic belief hierarchies in type space models

Definition 3.1. A Σ-based lexicographic type space (LT-space) is a tuple $(\Sigma, T_a, T_b, \tilde{g}_a, \tilde{g}_b)$, where T_i is a countable space endowed with the discrete topology and $\tilde{g}_i := (g_i^1, g_i^2, ...) : T_i \to \tilde{\Delta}(\Sigma \times T_j)$ maps every $t_i \in T_i$ to an LPS over $\Sigma \times T_j$.

[4]Our analysis can be generalized to any finite set of individuals.

An LT-space induces a hierarchy of lexicographic beliefs for every LT_i-type $t_i \in T_i$ as follows: The first order lexicographic beliefs are given by the LPS $\tilde{b}_i^1[t_i] \in \tilde{\Delta}(\Theta_0)$, where

$$b_i^{1,n}[t_i](\sigma) = \sum_{t_j \in T_j} g_i^n[t_i](\sigma, t_j), \text{ for all } \sigma \in \Theta_0 \text{ and all } n = 1, ..., N_{t_i}.$$

Let $\beta_j^1 : \tilde{\Delta}(\Theta_0) \to T_j \cup \{\emptyset\}$ associate every first order lexicographic beliefs to a type, whenever possible: $\beta_j^1(\tilde{p}_j^1) = \{t_j \in T_j : \tilde{b}_j^1[t_j] = \tilde{p}_j^1\}$ contains the types with first order lexicographic beliefs given by $\tilde{p}_j^1$. The second order lexicographic beliefs are given by the LPS $\tilde{b}_i^2[t_i] \in \tilde{\Delta}(\Theta_1)$ such that

$$\tilde{b}_i^{2,n}[t_i](\sigma, \tilde{p}_j^1) = \sum_{t_j \in \beta_j^1(\tilde{p}_j^1)} g_i^n[t_i](\sigma, t_j), \text{ for all } (\sigma, \tilde{p}_j^1) \in \Theta_1 \text{ and all } n = 1, ..., N_{t_i}.$$

Inductively, we define $\beta_j^k : \tilde{\Delta}(\Theta_{k-1}) \to T_j \cup \{\emptyset\}$, with $\beta_j^k(\tilde{p}_j^k) = \{t_j \in T_j : \tilde{b}_j^k[t_j] = \tilde{p}_j^k\}$. The k-th order lexicographic beliefs are given by the LPS $\tilde{b}_i^k[t_i] \in \tilde{\Delta}(\Theta_{k-1})$ such that

$$\tilde{b}_i^{k,n}[t_i](\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) = \sum_{t_j \in \bigcap_{l=1}^{k-1} \beta_j^l(\tilde{p}_j^l)} g_i^n[t_i](\sigma, t_j), \text{ for all } (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) \in \Theta_{k-1} \text{ and all } n = 1, ..., N_{t_i}.$$

The sequence $(\tilde{b}_i^1[t_i], \tilde{b}_i^2[t_i], ...)$ is t_i's lexicographic belief hierarchy. It is straightforward verifying that $(\tilde{b}_i^1[t_i], \tilde{b}_i^2[t_i], ...)$ is coherent for all $t_i \in T_i$, and every $i \in \{a, b\}$, implying that i weakly assumes j's coherency. However, i may still not assume that j's lexicographic belief hierarchies are coherent, as illustrated below.

Example 3.1. Let $\Sigma := \{\sigma_1, \sigma_2\}$ and $T_i := \{t_i^1, t_i^2, ...\}$ for each $i \in \{a, b\}$ endowed with the discrete topology. Let also $g_i^n[t_i]$ assign probability 1 to (σ_1, t_j^{2n}). Obviously, $\{(\sigma_2, t_j^1)\}$ is open in $\Sigma \times T_j$, while t_j^1 is associated with a coherent hierarchy. However, no theory of t_i assigns positive probability to $\{(\sigma, t_j^1)\}$, implying that i does not assume that j's beliefs are coherent. $\triangleleft$

We say that an LT_i-type satisfies mutual singularity whenever $\tilde{g}_i[t_i] \in \tilde{\Delta}(\Sigma \times T_j)$ is mutually singular. An LT-space satisfies mutual singularity whenever all all types of all individuals do. Unlike Brandenburger et al. (2008), we allow types with lexicographic beliefs that violate mutual singularity. We further discuss mutual singularity later in the paper.

3.2. Lexicographic belief hierarchies in standard state space models

Consider the measurable space $(\Omega, \mathcal{F})$, where Ω is Polish and $\mathcal{F}$ denotes the Borel σ-algebra. The continuous function $s : \Omega \to \Sigma$ determines the realized value of the underlying space of uncertainty at every state. Let $\mathcal{S}$ denote the partition of open subsets of Ω induced by s, with $S(\omega)$ being the element of $\mathcal{S}$ that contains ω. Let $\mathcal{P}_i$ be a partition of open subsets of Ω containing

i's (primarily) observable events: The set $P_i(\omega)$ is the element of $\mathcal{P}_i$ that contains ω, and denotes i's information set at ω, i.e., the states that i cannot distinguish from ω. This construction is due to Aumann (1976), and is often called Aumann space.

Lemma 3.1. *$\mathcal{P}_i$ is a countable partition.*

Definition 3.2. The tuple $(\Omega, \mathcal{F}, \mathcal{S}, \{\mathcal{P}_i\}_{i\in I}, \{\pi_i\}_{i\in I})$ is called a Σ-based standard state space (Ω-space), where $(\Omega, \mathcal{F})$ is a Polish space together with the Borel σ-algebra, $\mathcal{S}$ and $\mathcal{P}_i$ are partitions of open subsets of Ω, and π_i is a CPS on $(\Omega, \mathcal{F}, \mathcal{G}_i)$, where $\mathcal{P}_i \subseteq \mathcal{G}_i$.

Let $\mathcal{G}_i$ be constructed as follows: For $\omega \in \Omega$, we recursively define

$$P_i^n(\omega) := P_i^{n-1}(\omega) \setminus \Gamma\big(\pi_i(\cdot|P_i^{n-1}(\omega))\big), \tag{1}$$

with $P_i^1(\omega) := P_i(\omega)$, and let $\mathcal{G}_i = \bigcup_{\omega\in\Omega} \mathcal{G}_i^\omega$, with $\mathcal{G}_i^\omega := \{P_i^n(\omega)\ ;\ n = 1, ..., N_\omega\}$ being the collection of conditioning events that do not contradict the primarily observed hypothesis $P_i(\omega)$. It is straightforward verifying that $\mathcal{G}_i \subseteq \mathcal{F}$, which follows from $\Gamma\big(\pi_i(\cdot|P_i^n(\omega))\big)$ being closed.

Let the vector of probability measures $\pi_i(\cdot|\mathcal{G}_i^\omega) := \big(\pi_i(\cdot|P_i^n(\omega))\ ;\ n = 1, ..., N_\omega\big)$ be the collection of theories over $(\Omega, \mathcal{F})$ corresponding to the different conditioning events in $\mathcal{G}_i^\omega$. The fact that $\pi_i(\cdot|\mathcal{G}_i^\omega)$ is an ordered collection of measures, implies that it is an LPS over $(\Omega, \mathcal{F})$, which furthermore induces a lexicographic belief hierarchy at ω.

Construct the following LPS over $\Theta_0 = \Sigma$: Let $\tilde{\mu}_{i,\omega}^1 \in \tilde{\Delta}(\Theta_0)$ be i's first order lexicographic beliefs at ω, where

$$\mu_{i,\omega}^{1,n}(\sigma) := \pi_i(\{\omega' : s(\omega') = \sigma\}|P_i^n(\omega)), \text{ for all } \sigma \in \Theta_0 \text{ and all } n = 1, ..., N_\omega.$$

Obviously, $\{\omega' : s(\omega') = \sigma\}$ is $\mathcal{B}_j^0$-measurable, where $\mathcal{B}_j^0 := \mathcal{S}$, implying that $\mu_{i,\omega}^{1,n}(\sigma)$ is a well-defined probability.

Let the partition $\mathcal{B}_j^1$ be the collection of j's first order lexicographic beliefs equivalence classes, i.e., two states ω and ω' belong to the same element of $\mathcal{B}_j^1$ if and only if $\tilde{\mu}_{j,\omega}^1 = \tilde{\mu}_{j,\omega'}^1$. Construct the following LPS over $\Theta_1 = \Theta_0 \times \tilde{\Delta}(\Theta_0)$: Let $\tilde{\mu}_{i,\omega}^2 \in \tilde{\Delta}(\Theta_1)$ be i's second order lexicographic beliefs at ω, where

$$\mu_{i,\omega}^{2,n}(\sigma, \tilde{p}_j^1) := \pi_i(\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1) = (\sigma, \tilde{p}_j^1)\}|P_i^n(\omega)), \text{ for all } (\sigma, \tilde{p}_j^1) \in \Theta_1 \text{ and all } n = 1, ..., N_\omega.$$

The set $\{\omega' : s(\omega') = \sigma\}$ is $\mathcal{B}_j^0$-measurable, while $\{\omega' : \tilde{\mu}_{j,\omega'}^1 = \tilde{p}_j^1\}$ is $\mathcal{B}_j^1$-measurable, implying that $\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1) = (\sigma, \tilde{p}_j^1)\}$ is $(\mathcal{B}_j^0 \vee \mathcal{B}_j^1)$-measurable. Therefore, $\mu_{i,\omega}^{2,n}(\sigma, \tilde{p}_j^1)$ is a well-defined probability.

Likewise, inductively define the partition $\mathcal{B}_j^k$, and construct the following LPS over $\Theta_{k-1} = \Theta_0 \times \tilde{\Delta}(\Theta_1) \times \cdots \times \tilde{\Delta}(\Theta_{k-2})$: Let $\tilde{\mu}_{i,\omega}^k$ be i's k-th order lexicographic beliefs over Θ_{k-1} at ω, where

$$\begin{aligned} \mu_{i,\omega}^{k,n}(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) \quad := \quad & \pi_i(\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1 ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1 ..., \tilde{p}_j^{k-1})\} | P_i^n(\omega)), \\ & \text{for all } (\sigma, \tilde{p}_j^1 ..., \tilde{p}_j^{k-1}) \in \Theta_{k-1} \text{ and all } n = 1, ..., N_\omega. \end{aligned}$$

Finally, it is straightforward verifying that $\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1 ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1 ..., \tilde{p}_j^{k-1})\}$ is $(\mathcal{B}_j^0 \vee \cdots \vee \mathcal{B}_j^{k-1})$-measurable. Therefore, $\mu_{i,\omega}^{k,n}(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})$ is a well-defined probability.

The sequence $(\tilde{\mu}_{i,\omega}^1, \tilde{\mu}_{i,\omega}^2, ...)$ is i's lexicographic belief hierarchy at ω. Let $\mathcal{B}_i := \bigvee_{k=1}^{\infty} \mathcal{B}_i^k$ denote the coarsest common refinement (join) of the partitions $\mathcal{B}_i^k$. It follows from $\mathcal{B}_i$ being itself a coarsening of $\mathcal{P}_i$ that i has the same lexicographic belief hierarchy at all states in $P_i(\omega)$. Therefore, $P_i(\omega)$ can be treated as one element, henceforth called i's type in the Ω-space ($\mathcal{P}_i$-type).

Since i forms a hierarchy of lexicographic beliefs at every state ω, coherency can be expressed as an event in Ω. Likewise, weak assumption of coherency is an event. Therefore, we can iteratively define common weak assumption of coherency in the Ω-space.

Proposition 3.1. $(\tilde{\mu}_{i,\omega}^1, \tilde{\mu}_{i,\omega}^2, ...)$ *satisfies common weak assumption of coherency.*

Similarly to the case illustrated in Example 3.1, the previous result cannot be extended to common assumption of coherency. The event C_j, containing the states where j's lexicographic beliefs are coherent, is assumed by i at ω, whenever for every open $F \subseteq P_i(\omega)$ there is some $i = 1, ..., N_\omega$ such that $\pi_i(F \cap C_j | P_i^n(\omega)) > 0$. Obviously, if the open set $P_i(\omega) \cap P_j(\omega) \neq \emptyset$ is not assigned positive probability by any theory of i at ω, then i does not assume j's coherency. The following example illustrates such a situation.

Example 3.2. Let $\Sigma := \{\sigma_1, \sigma_2\}$ and $\Omega_i := \{\omega_i^1, \omega_i^2, ...\}$ for each $i \in \{a, b\}$, and consider the space $\Omega := \Sigma \times \Omega_a \times \Omega_b$ endowed with the discrete topology. Define $\mathcal{P}_i$ as follows: For each $(\sigma, \omega_i, \omega_j) \in \Omega$, let $P_i(\sigma, \omega_i, \omega_j) := \{(\sigma', \omega_i', \omega_j') \in \Omega : \omega_i' = \omega_i\}$. Let $\pi_i(\cdot | P_i^n(\sigma, \omega_i, \omega_j))$ assign probability 1 to $(\sigma_1, \omega_i, \omega_j^{2n})$, implying that $P_j(\sigma, \omega_i, \omega_j^1)$ is assigned zero probability by all theories of i at $(\sigma, \omega_i, \omega_j)$, and therefore i does not assume that j's lexicographic beliefs are coherent. $\triangleleft$

3.3. Epistemic equivalence

The notion of epistemic equivalence between Ω-spaces and LT-spaces requires that

(i) there is a bijection between $\mathcal{P}_i$-types and LT_i-types, and

(ii) the corresponding types are associated with the same lexicographic belief hierarchy.

3.3.1. From state spaces to type spaces

In this section, we show that for every Ω-space there is an associated epistemically equivalent LT-space, constructed as follows: We consider the set of $\mathcal{P}_i$-types, and map every information set in $\mathcal{P}_i$ to an LPS over $(\Omega, \sigma(\mathcal{S} \vee \mathcal{P}_j))$. Finally, we show that this construction is a Σ-based LT-space, which in addition is epistemically equivalent to the original Ω-space.

Formally, let T_i, endowed with the discrete topology, be bijective to $\mathcal{P}_i$, and consider the natural mapping

$$f_i : \Omega \to T_i, \tag{2}$$

such that $f_i(\omega) = f_i(\omega')$ if and only if $P_i(\omega) = P_i(\omega')$.

Proposition 3.2. *Consider an Ω-space and the function $f_i : \Omega \to T_i$. Then, there is an epistemically equivalent LT-space $(\Sigma, T_a, T_b, \tilde{g}_a, \tilde{g}_b)$, i.e., for all $\omega \in \Omega$,*

$$(\tilde{\mu}^1_{i,\omega}, \tilde{\mu}^2_{i,\omega}, ...) = (\tilde{b}^1_i[f_i(\omega)], \tilde{b}^2_i[f_i(\omega)], ...).$$

3.3.2. From type spaces to state spaces

Let $W := \Sigma \times T_a \times T_b$, and for every $w \in W$, let Ω_w be a countable space endowed with the discrete topology, with typical element ω_w. It follows from Ω_w being Polish, that

$$\Omega_\infty := \oplus_{w \in W} \Omega_w$$

is Polish too (Srivastava, 1991, p. 52). Moreover, it follows from the definition of the topological sum[5] that every Ω_w is open in Ω_∞.

Consider the natural mapping $f_i : \Omega_\infty \to T_i$ defined by $f_i(\omega_w) := \text{proj}_{T_i} w$, where proj denotes the projection. Let $\mathcal{P}_i$ be partition of type equivalent classes in Ω_∞, with

$$P_i(\omega) := \{\omega' \in \Omega_\infty : f_i(\omega') = f_i(\omega)\}.$$

Likewise, let $s : \Omega_\infty \to \Sigma$ be defined as $s(\omega_w) := \text{proj}_\Sigma w$, and consider the partition $\mathcal{S}$ of Σ-realization-equivalent classes, i.e.,

$$S(\omega) := \{\omega' \in \Omega_\infty : s(\omega') = s(\omega)\}.$$

It follows from $\mathcal{S}$ and $\mathcal{P}_i$ being coarsenings of $\{\Omega_w \,;\, w \in W\}$, that they are both partitions of open subsets of Ω_∞.

[5] An event $E \subseteq \oplus_{w \in W} \Omega_w$ is open in Ω_∞ if and only if $E \cap \Omega_w$ is open in Ω_w for all $w \in W$.

Proposition 3.3. *Consider an LT-space and the Polish space Ω_∞ together with the (discrete) Borel σ-algebra $\mathcal{F}$. Let the natural mappings $f_i : \Omega_\infty \to T_i$ and $s : \Omega_\infty \to \Sigma$ induce the partitions of open sets $\mathcal{P}_i$ and $\mathcal{S}$ respectively. Then, there is an epistemically equivalent Ω-space $(\Omega_\infty, \mathcal{F}, \{\mathcal{G}_i\}_{i\in I}, \mathcal{S}, \{\mathcal{P}_i\}_{i\in I}, \{\pi_i\}_{i\in I})$, i.e., for all $\omega \in \Omega_\infty$,*

$$(\tilde{\mu}^1_{i,\omega}, \tilde{\mu}^2_{i,\omega}, ...) = (\tilde{b}^1_i[f_i(\omega)], \tilde{b}^2_i[f_i(\omega)], ...).$$

In their seminal paper, Brandenburger and Dekel (1993) construct the canonical state space of the universal belief space, as the product space $\Sigma \times T_a \times T_b$, implying that in their setting every Ω_w is a singleton. The reason for adopting a richer state space is in order to be able to map *LT*-spaces that violate mutual singularity to Ω-spaces with conditioning events satisfying $P^n_i(\omega) := P^{n-1}_i(\omega) \setminus \Gamma\big(\pi_i(\cdot|P^{n-1}_i(\omega))\big)$. That would not have been possible if we had assumed that all Ω_w were singletons. We further discuss this issue in the following section.

4. Discussion

4.1. Mutual singularity

Recall the definition of a mutually singular LPS: We say that $(p^n)^N_{n=1}$ satisfies mutual singularity in Z whenever for each $n = 1, ..., N$ there is some Borel set F_n such that $p^n(F_n) = 1$ and $p^n(F_m) = 0$ for all $m \neq n$. An *LT*-space is mutually singular whenever, for all $t_i \in T_i$ the LPS $\tilde{g}_i[t_i]$ satisfies mutual singularity in $\Sigma \times T_j$ (Brandenburger et al., 2008).

Mutual singularity aims at capturing the idea that the different conditioning hypotheses, which are associated with the different theories, overlap – roughly speaking – as little as possible. However, in an *LT*-space the different conditioning events are not clearly specified.

On the other hand, in an Ω-space, conditioning hypotheses are clearly defined in a very natural way (see Eq. (1)). We say that an Ω-space satisfies mutual singularity, whenever the epistemically equivalent *LT*-space (following from Proposition 3.2) is mutually singular. At a first glance, this last definition seems to potentially contradict the fact that $\pi_i(\cdot|\mathcal{G}^\omega_i)$ is an LPS satisfying mutual singularity. However, if we take a closer look, it becomes clear that no such contradiction exists, as $\pi_i(\cdot|\mathcal{G}^\omega_i)$ is an LPS over $(\Omega, \mathcal{F})$, whereas mutual singularity is defined over $\Sigma \times T_j$, which corresponds to the measurable space $(\Omega, \sigma(\mathcal{S} \vee \mathcal{P}_j))$. The following result provides sufficient conditions for a mutually singular Ω-space. Let $\mathcal{J} := \mathcal{S} \vee \mathcal{P}_a \vee \mathcal{P}_b$ be the coarsest common refinement of the three partitions.

Proposition 4.1. *If $P^n_i(\omega)$ is $\sigma(\mathcal{J})$-measurable at each $\omega \in \Omega$, for all $i \in \{a, b\}$, and all $n > 0$, then the Ω-space satisfies mutual singularity.*

The previous proposition also explains why taking every Ω_w as a singleton does not suffice for mapping an LT-space that violates mutual singularity to an Ω-space: It follows directly from Proposition 4.1 that if the corresponding LT-space is not mutually singular then there is some $P_i^n(\omega)$ which is not $\sigma(\mathcal{J})$-measurable, which cannot be the case when every Ω_w is a singleton. This follows from the fact that $\mathcal{J} = \{\Omega_w \,;\, w \in W\}$.

4.2. Redundancies

Recall the standard definition of redundancy by Mertens and Zamir (1985, p. 6), generalized from usual to lexicographic beliefs: An LT-space is redundant if there are two distinct types of the same individual associated with the same belief hierarchy, i.e., there are $t_i, t_i' \in T_i$ such that $(\tilde{b}_i^1[t_i], \tilde{b}_i^2[t_i], ...) = (\tilde{b}_i^1[t_i'], \tilde{b}_i^2[t_i'], ...)$. Friedenberg and Meier (2010) introduced a weaker form of redundancy: An LT-space is measurably non-redundant whenever all types inducing the same lexicographic belief hierarchy cannot be separated by measurable sets, and it is redundant otherwise, i.e., i's beliefs are measurably non-redundant whenever $(\tilde{b}_i^1[t_i], \tilde{b}_i^2[t_i], ...) = (\tilde{b}_i^1[t_i'], \tilde{b}_i^2[t_i'], ...)$ implies that for every measurable subset E of T_i, either $\{t_i, t_i'\} \subseteq E$ or $\{t_i, t_i'\} \subseteq T_i \setminus E$. The latter restricts attention only to redundancies which can be expressed with the language induced by the corresponding σ-algebra on T_i. Obviously, if the language is fine enough to separate every type, like for instance in the LT-spaces we consider where T_i is endowed with the discrete topology, redundancy and measurable redundancy coincide. Henceforth, for the purposes of this paper, we deem the two notions identical.

Extending the definition of redundancy to Ω-spaces is straightforward: An Ω-space is redundant if there are two $\mathcal{P}_i$-types inducing the same lexicographic belief hierarchy, i.e., there are $\omega, \omega' \in \Omega$ with $P_i(\omega) \cap P_i(\omega') = \emptyset$ such that $(\tilde{\mu}_{i,\omega}^1, \tilde{\mu}_{i,\omega}^2, ...) = (\tilde{\mu}_{i,\omega'}^1, \tilde{\mu}_{i,\omega'}^2, ...)$. The following result characterizes redundancy in an Ω-space. Recall that $\mathcal{B}_i := \bigvee_{k=1}^{\infty} \mathcal{B}_i^k$, where $\mathcal{B}_i^k$ is the collection of k-th order lexicographic belief equivalence classes, i.e., two states belong to the same element of $\mathcal{B}_i^k$ if and only if they induce the same k-th order lexicographic beliefs.

Proposition 4.2. *An Ω-space is non-redundant if and only if $\mathcal{B}_i = \mathcal{P}_i$.*

It is straightforward verifying that the equivalences proven in Propositions 3.2 and 3.3 preserve the belief redundancies, as the bijection between the types preserves the lexicographic belief hierarchies.

4.3. Hierarchies of conditional beliefs

Battigalli and Siniscalchi (1999) defined the notion of a hierarchy of conditional beliefs, which is

conceptually very similar to a lexicographic belief hierarchy. Individuals start with a common collection of non-empty conditioning hypotheses $\mathcal{G}$. A hierarchy of conditional beliefs is a sequence of conditional probability systems given the conditioning events: The first order conditional beliefs is a collection of measures, one for every $G \in \mathcal{G}$, over the underlying space of uncertainty; the second order beliefs is a collection of measures, one for every $G \in \mathcal{G}$, over the product of the underlying space of uncertainty and the space of the opponent's first order conditional beliefs; likewise the entire hierarchy is constructed.

Like most types of belief hierarchies, hierarchies of conditional beliefs can be represented by a type space model (Battigalli and Siniscalchi, 1999): Each type is endowed with a CPS over the product of the fundamental space of uncertainty and the opponent's set of types. This construction associates every type to a hierarchy of conditional beliefs.

Similarly to lexicographic belief hierarchies, hierarchies of conditional beliefs represent probabilistic assessments given a collection of conditioning hypotheses. The main difference between the two approaches is that in the former different individuals condition with respect to different events, whereas hierarchies of conditional beliefs restrict to a common collection of hypotheses. In any case, the LPS $\pi_i(\cdot|\mathcal{G}_i^\omega)$ associated with the type $P_i(\omega)$ in $(\Omega, \mathcal{F})$, is also a CPS over $(\Omega, \mathcal{F}, \mathcal{G}_i^\omega)$, and therefore if we allow for different conditioning events, the lexicographic belief hierarchy becomes a hierarchy of conditional beliefs.

4.4. Uncountable spaces

Our analysis, for the sake of presentation simplicity, is restricted to countable type spaces. We could generalize most of our results to arbitrary structures, in which case we would need to impose some additional topological assumptions, e.g., we would need to substitute information partitions with σ-algebras (see, Brandenburger and Dekel, 1993, p. 196).

Appendix

Proof of Lemma 3.1. Since Ω is Polish, it is second countable, implying that it has a countable basis. It follows from $P_i(\omega)$ being open that it can be written as the union of elements of this basis. Since $\mathcal{P}_i$ partitions Ω, it follows that the elements of the basis used to generated $P_i(\omega)$ are distinct to those used to generate $P_i(\omega')$, whenever $P_i(\omega) \cap P_i(\omega') = \emptyset$. Therefore, the cardinality of $\mathcal{P}_i$ at most equal to the cardinality of the basis, implying that $\mathcal{P}_i$ is countable. □

Proof of Proposition 3.1. It suffices to show that $(\tilde{\mu}_{i,\omega}^1, \tilde{\mu}_{i,\omega}^2, ...)$ satisfies common weak assumption of coherency for all $\omega \in \Omega$ and all $i \in I$. For any $k > 1$ and $n > 0$, and given an arbitrary $(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-2}) \in \Theta_{k-2}$,

we obtain

$$\begin{aligned}
\sum_{\tilde{p}_j^{k-1}\in\tilde{\Delta}(\Theta_{k-2})} \mu_{i,\omega}^{k,n}(\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-1}) &= \sum_{\tilde{p}_j^{k-1}\in\tilde{\Delta}(\Theta_{k-2})} \pi_i(\{\omega' : (s(\omega'),\tilde{\mu}_{j,\omega'}^1...,\tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma,\tilde{p}_j^1...,\tilde{p}_j^{k-1})\}|P_i^n(\omega)) \\
&= \pi_i(\{\omega' : (s(\omega'),\tilde{\mu}_{j,\omega'}^1...,\tilde{\mu}_{j,\omega'}^{k-2}) = (\sigma,\tilde{p}_j^1...,\tilde{p}_j^{k-2})\}|P_i^n(\omega)) \\
&= \mu_{i,\omega}^{k-1,n}(\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-2}),
\end{aligned}$$

implying $\text{marg}_{\Theta_{k-2}}\,\mu_{i,\omega}^{k,n} = \mu_{i,\omega}^{k-1,n}$, and therefore $\text{marg}_{\Theta_{k-2}}\,\tilde{\mu}_{i,\omega}^{k} = \tilde{\mu}_{i,\omega}^{k-1}$. □

Proof of Proposition 3.2. Define $g_i : T_i \to \tilde{\Delta}(\Sigma \times T_j)$ as follows,

$$g_i^n[f_i(\omega)](\sigma,t_j) := \pi_i\big(\{\omega' : s(\omega') = \sigma\} \cap \{\omega' : f_j(\omega') = t_j\}|P_i^n(\omega)\big),$$

which is a well-defined probability as f_j is continuous and therefore $\{\omega' : f_j(\omega') = t_j\}$ is Borel. For any $\omega \in \Omega$, $k > 0$ and $n > 0$, take an arbitrary $(\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-1}) \in \tilde{\Delta}(\Theta_{k-1})$. Then,

$$\begin{aligned}
b_i^{k,n}[f_i(\omega)](\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-1}) &= \sum_{t_j\in\bigcap_{l=1}^{k-1}\beta_j^l(\tilde{p}_j^l)} g_i^n[f_i(\omega)](\sigma,t_j) \\
&= \sum_{t_j\in\bigcap_{l=1}^{k-1}\beta_j^l(\tilde{p}_j^l)} \pi_i\big(\{\omega' : s(\omega') = \sigma\} \cap \{\omega' : f_j(\omega') = t_j\}|P_i^n(\omega)\big) \\
&= \pi_i\big(\{\omega' : (s(\omega'),\tilde{\mu}_{j,\omega'}^1,...,\tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-1})\}|P_i^n(\omega)\big) \\
&= \mu_{i,\omega}^{k,n}(\sigma,\tilde{p}_j^1,...,\tilde{p}_j^{k-1})
\end{aligned}$$

completes the proof. □

Proof of Proposition 3.3. For each $w = (\sigma,t_i,t_j) \in W$, let $\Omega_w := \{\omega_w^1,\omega_w^2,...\}$. Let also $\mathcal{I}_{w,i} := \{I_{w,i}^1,...,I_{w,i}^{N_{t_i}}\}$ be a sequence of subsets of Ω_w, such that every $I_{w,i} \in \mathcal{I}_{w,i}$ contains at most one ω_w^n with $g_i^n[t_i](\sigma,t_j) > 0$, and is empty whenever $g_i^n[t_i](\sigma,t_j) = 0$. Moreover, $\mathcal{I}_{w,i}$ covers Ω_w. That is, roughly speaking, we construct $\mathcal{I}_{w,i}$ as follows: Take the natural numbers satisfying $g_i^n[t_i](\sigma,t_j) > 0$ and include them in the corresponding $I_{w,i}^n$. Allocate the remaining states into these non-empty $I_{w,i}^n \in \mathcal{I}_{w,i}$. The elements of $\mathcal{I}_{w,i}$ that have not been filled with any state will remain empty. Obviously, if $g_i^n[t_i](\sigma,t_j) = 0$ for all $n = 1,...,N_{t_i}$, then all $I_{w,i}^n \in \mathcal{I}_{w,i}$ remain empty. Formally, $\mathcal{I}_{w,i}$ satisfies

(i) $I_{w,i}^n \neq \emptyset$ if and only if $g_i^n[t_i](\sigma,t_j) > 0$,

(ii) $I_{w,i}^n \cap I_{w,i}^m \neq \emptyset$ if $g_i^n[t_i](\sigma,t_j) > 0$ and $g_i^m[t_i](\sigma,t_j) > 0$, and

(iii) $\bigcup_{n=1}^{N_{t_i}} I_{w,i}^n = \Omega_w$ if $\Gamma(\tilde{g}_i[t_i]) \neq \emptyset$.

By construction $\mathcal{I}_{w,i}$ always exists. Define i's sequence of conditioning events at $w = (\sigma,t_i,t_j)$ as follows:

$$P_i^n(\omega) := P_i^{n-1}(\omega) \setminus \Big(\bigcup_{w'\in W\,:\,\text{proj}_{T_i}\,w'=t_i} I_{w',i}^{n-1}\Big),$$

with $P_i^1(\omega) = P_i(\omega)$. Finally, as usual, let $\mathcal{G}_i^\omega := \{P_i^n(\omega)\ ;\ n > 0\}$ and $\mathcal{G}_i := \bigcup_{\omega\in\Omega_\infty}\mathcal{G}_i^\omega$.

We define the CPS π_i over $(\Omega_\infty, \mathcal{F}, \mathcal{G}_i)$ as follows: For an arbitrary $\omega \in \Omega_\infty$, and for all $w = (\sigma, t_i, t_j) \in W$, let

$$\pi_i(\Omega_w | P_i^n(\omega)) = \begin{cases} g_i^n[t_i](\sigma, t_j) & \text{if } \Omega_w \subseteq P_i(\omega) \\ 0 & \text{if } \Omega_w \cap P_i(\omega) = \emptyset \end{cases}$$

Moreover, let $\pi_i(\cdot | P_i^n(\omega) \cap \Omega_w)$ be concentrated, e.g., uniformly distributed, on states in $I_{w,i}^n$. Obviously, if $I_{w,i}^n = \emptyset$, which is the case when $g_i^n[t_i](\sigma, t_j) = 0$ (see (i) above), then $\pi_i(\cdot | P_i^n(\omega) \cap \Omega_w)$ is "concentrated" on the empty set, implying $\pi_i(F | P_i^n(\omega) \cap \Omega_w) = 0$ for all $F \in \mathcal{F}$. Thus, let

$$\pi_i(\cdot | P_i^n(\omega)) := \sum_{\Omega_w \subseteq P_i(\omega)} \pi_i(\cdot | P_i^n(\omega) \cap \Omega_w) \pi_i(\Omega_w | P_i^n(\omega)).$$

It is straightforward verifying that $P_i^n(\omega) = P_i^{n-1}(\omega) \setminus \Gamma\big(\pi_i(\cdot | P_i^{n-1}(\omega))\big)$.

Consider some $(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) \in \tilde{\Delta}(\Theta_{k-1})$, and for each $k > 0$, and every $n = 1, ..., T_{t_i}$,

$$\begin{aligned} \mu_{i,\omega}^{k,n}(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) &= \pi_i\big(\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1, ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})\} | P_i^n(\omega)\big) \\ &= \sum_{\Omega_w \subseteq P_i(\omega)} \pi_i(\Omega_w | P_i^n(\omega)) \times \\ &\quad \times \pi_i(\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1, ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})\} | P_i^n(\omega) \cap \Omega_w). \end{aligned} \tag{3}$$

Observe that $\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1, ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})\}$ is $\sigma(\{\Omega_w \,;\, w \in W\})$-measurable, implying that $\pi_i(\{\omega' : (s(\omega'), \tilde{\mu}_{j,\omega'}^1, ..., \tilde{\mu}_{j,\omega'}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})\} | P_i^n(\omega) \cap \Omega_w) \in \{0, 1\}$, where it is equal to 1 if and only if ω_w satisfies

(a) $(s(\omega_w), \tilde{\mu}_{j,\omega_w}^1, ..., \tilde{\mu}_{j,\omega_w}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})$, and

(b) $g_i^n[f_i(\omega_w)](\sigma, t_j) > 0$.

Therefore, we rewrite (4) as follows:

$$\begin{aligned} \mu_{i,\omega}^{k,n}(\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}) &= \sum_{w \in W : (s(\omega_w), \tilde{\mu}_{j,\omega_w}^1, ..., \tilde{\mu}_{j,\omega_w}^{k-1}) = (\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1})} \pi_i(\Omega_w | P_i^n(\omega)) \\ &= \sum_{t_j \in \bigcap_{l=1}^{k-1} \beta_j^l(\tilde{p}_j^l)} g_i^n[f_i(\omega)](\sigma, t_j) \\ &= b_i^{k,n}[f_i(\omega)](\sigma, \tilde{p}_j^1, ..., \tilde{p}_j^{k-1}), \end{aligned}$$

which completes the proof. □

Proof of Proposition 4.1. It follows by induction that, if $P_i^n(\omega) \in \sigma(\mathcal{J})$ for all $n > 0$, then $\Gamma\big(\pi_i(\cdot | P_i^n(\omega))\big) \in \sigma(\mathcal{J})$ for all $n > 0$, implying that for every $J \in \mathcal{J}$,

$$\text{either } J \subseteq \Gamma\big(\pi_i(\cdot | P_i^n(\omega))\big) \text{ or } \Gamma\big(\pi_i(\cdot | P_i^n(\omega))\big) \cap J = \emptyset. \tag{4}$$

For some $\omega \in \Omega$, define the sequence of Borel sets in $\Sigma \times T_j$,

$$F_{i,\omega}^n := \{\ (\sigma, t_j)\ :\ \{\omega' : s(\omega') = \sigma\} \cap \{\omega' : f_j(\omega') = t_j\} \cap \{\omega' : f_i(\omega') = f_i(\omega)\} \subseteq \Gamma\big(\pi_i(\cdot | P_i^n(\omega))\big)\ \}\ .$$

It follows from (4) and $\{\omega' : s(\omega') = \sigma\} \cap \{\omega' : f_j(\omega') = t_j\} \cap \{\omega' : f_i(\omega') = f_i(\omega)\}$ being $\sigma(\mathcal{J})$-measurable that for all $m > 0$,

$$g_i^n[f_i(\omega)](F_{i,\omega}^m) = \pi_i\Big(\Gamma\big(\pi_i(\cdot|P_i^m(\omega))\big) \;\Big|\; P_i^n(\omega)\Big).$$

Consider the following cases:

(i) If $m = n$, then it follows from the definition of the support that $g_i^n[f_i(\omega)](F_{i,\omega}^n) = 1$.

(ii) If $m < n$, then it follows from the definition of the conditioning events that $\Gamma\big(\pi_i(\cdot|P_i^m(\omega))\big) \cap P_i^n(\omega) = \emptyset$, implying that $g_i^n[f_i(\omega)](F_{i,\omega}^m) = 0$.

(iii) Let $m > n$. It follows from Lemma 3.1 that $\mathcal{J}$ is a partition of open subsets, and therefore every $\sigma(\mathcal{J})$-measurable event is closed. Hence, $P_i^m(\omega)$ is closed. Therefore, it follows, from $\Gamma\big(\pi_i(\cdot|P_i^m(\omega))\big) \in \sigma(\mathcal{J})$, that $\Gamma\big(\pi_i(\cdot|P_i^m(\omega))\big) \subseteq P_i^m(\omega)$, implying

$$\pi_i\Big(\Gamma\big(\pi_i(\cdot|P_i^m(\omega))\big) \;\Big|\; P_i^n(\omega)\Big) \leq \pi_i\big(P_i^m(\omega)|P_i^n(\omega)\big) = 0.$$

That is, $g_i^n[f_i(\omega)](F_{i,\omega}^m) = 0$.

It follows from (i)–(iii) that $\tilde{g}_i[f_i(\omega)]$ satisfies mutual singularity, which completes the proof. □

Proof of Proposition 4.2. It follows by construction that $\mathcal{B}_i$ is a partition, weakly coarser than $\mathcal{P}_i$. First, let $\mathcal{P}_i = \mathcal{B}_i$, and consider $\omega, \omega' \in \Omega$ such that $P_i(\omega) \cap P_i(\omega') = \emptyset$. Then, it follows directly that ω and ω' belong to different elements of $\mathcal{B}_i$, implying that they induce different lexicographic belief hierarchies, and therefore Ω is non-redundant. For the converse, suppose that Ω is non-redundant, which implies that every $\mathcal{P}_i$-type yields a different belief hierarchy, which completes the proof. □

References

ALIPRANTIS, C. & BORDER, K. (1994). *Infinite dimensional analysis*. Springer Verlag, Berlin.

AUMANN, R.J. (1976). Agreeing to disagree. *Annals of Statistics* 4, 1236-1239.

——– (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55, 1–18.

AUMANN, R.J. & BRANDENBURGER, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica* 63, 1161–1180.

BATTIGALLI, P. & SINISCALCHI, M. (1999). Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory* 88, 188–230.

——— (2002). Strong belief and forward induction reasoning. *Journal of Economic Theory* 106, 356–391.

BERNHEIM, D. (1984). Rationalizable strategic behavior. *Econometrica* 52, 1007–1028.

BLACKWELL, D. & DUBINS, L. (1975). On existence and non-existence of proper, regular, conditional distributions. *Annals of Probability* 3, 741–752.

BLACKWELL, D. & YLL-NARDZEWSKI, C. (1963). Non-existence of everywhere proper conditional distributions. *Annals of Mathematical Statistics* 34, 223–225.

BLUME, L., BRANDENBURGER, A. & DEKEL, E. (1991a). Lexicographic probabilities and choice under uncertainty. *Econometrica* 59, 61–79.

——— (1991b). Lexicographic probabilities and equilibrium refinements. *Econometrica* 59, 81–98.

BRANDENBURGER, A. & DEKEL, E. (1987). Common knowledge with probability 1. *Journal of Mathematical Economics* 16, 237–245.

——— (1987). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory* 59, 189–198.

BRANDENBURGER, A. & FRIEDENBERG, A. (2010). Self-admissible sets. *Journal of Economic Theory* 145, 785–811.

BRANDENBURGER, A., FRIEDENBERG, A. & KEISSLER, H.J. (2007). Notes on the relationship between strong belief and assumption. Unpublished manuscript.

——— (2008). Admissibility in games. *Econometrica* 76, 307–352.

FRIEDENBERG, A. & MEIER, M. (2010). On the relationship between hierarchy and type morphisms. *Economic Theory*, forthcoming.

HALPERN, J.Y. (2010). Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior* 68, 155–179.

HARSANYI, J. (1967-68). Games with incomplete information played by Bayesian players, I-III. *Management Science* 14, 159–182, 320–334, 486–502.

KECHRIS, A. (1995). *Classical descriptive set theory.* Springer Verlag, Berlin.

MERTENS, J.F. & ZAMIR, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory* 14, 1–29.

PARTHASARATHY, K.R. (1967). *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence, Rhode Island.

PEARCE, D.G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52, 1029–1050.

PEREA, A. (2010). *Reasoning and choice: An epistemic course in game theory.* Cambridge University Press, forthcoming.

RÊNYI, A. (1955). On a new axiomatic theory of probability. *Acta Mathetica Academiae Scientiarum Hungaricae* 6, 285–335.

SELTEN, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive form games. *International Journal of Game Theory* 4, 25–55.

SRIVASTAVA, S.M. (1991). *A course on Borel sets.* Springer Verlag, Berlin.

TAN, T. & WERLANG, S. (1992). On Aumann's notion of common knowledge: An alternative approach. *Revista Brasileira de Economia* 64, 151–166.